Electronic Supplementary Material (ESI) for Chemical Science.

Exploring modular reengineering strategies to redesign the teicoplanin non-ribosomal peptide synthetase

Milda Kaniusaite,[a,b] Robert J. A. Goode,[a,c] Julien Tailhades,[a,b] Ralf B. Schittenhelm[a,c] and Max J. Cryle*[a,b]

a. The Monash Biomedicine Discovery Institute, Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria 3800, Australia.

b. EMBL Australia, Monash University, Clayton, Victoria 3800, Australia.

c. Monash Proteomics and Metabolomics Facility, Monash University, Clayton, Victoria 3800, Australia.

Supplementary Information

Supplementary Table S1. Primer sequences and template DNA used for construct generation

Final construct	Primers	Insert template DNA	Vector template DNA
M5-6 (A-PCP-E)$_5$-(C-A-PCP)$_6$	Fr-1-Mod5_ATEC.FOR AATCTTTATTTTCAGGGGAGCAGCGTA F1_Mod5_ATEC.REV GGTTGTGCCGTGAGCGGCGCATCTGTT TCGTTCC	pET GB1 1d Tcp11 module 5 (ATEC)	-
	Fr2_Mod6_AT.REV CCAAGCACTCTCGAGTTCACCGGCCGG F2_Mod6_AT.FOR GCTCACGGCACAACCGCTTCCCAT	pET GB1 1d Tcp11 module 6 (AT)	-
	Vector56_ATECAT.REV CTGAAAATAAAGATTCTCAGAACCACT GCCA Vector56_ATECAT.FOR CTCGAGAGTGCTTGGAGTCATCC	-	pET GB1 1d series
M5-6a (A-PCP-E)$_5$-(C-A-PCP)$_6$-E$_2$	Fr1-56.FOR AATCTTTATTTTCAGGGGAGCAGCGTA Fr1-56.REV GCCACTTGCGCCGCCAGACGTTCC	pET GB1 1d Tcp11 dimodule 5-6 (ATEC-AT)	-
	Fr2-Tcp9.FOR GGCGGCGCAAGTGGCAGCGCA Fr2-Tcp9.REV GCAGCCGGATCAAGCTTACTTCTCGAA CTGTGGATGACTCCAAG	pET SUMO 1d Tcp9	-
	Vector-56E.REV	-	pET GB1 1d Tcp11

Final construct	**Primers**	**Insert template DNA**	**Vector template DNA**
	CTGAAAATAAAGATTCTCAGAACCACT GCCA Vector-56E.FOR GCTTGATCCGGCTGCTAACAAAG		dimodule 5-6 (ATEC-AT)
M5-6b $(A\text{-}PCP\text{-}E)_5\text{-}(C\text{-}A)_6\text{-}(PCP\text{-}E)_2$	F1-56.FOR AGAAGGAGATATACCATGAAACATCAC CATC F1-56.REV ACTTCCGTCCGAGGATCACGCCC	pET GB1 1d Tcp11 dimodule 5-6 (ATEC-AT)	-
	F2-PCP-E(tcp9).FOR TCCTCGGACGGAAGTCGAGCGCGC F2-PCP-E(tcp9).REV GCAGCCGGATCAAGCTTACTTCTCGAA CTGTGGATGACTCCAAG	pET SUMO 1d Tcp9	-
	V-56_PCP-E.REV GGTATATCTCCTTCTTAAAGTTAAACAA AATTATTTCTAGAGGG V-56_PCP-E.FOR GCTTGATCCGGCTGCTAACAAAG	-	pET GB1 1d 5-6_E $(A\text{-}PCP\text{-}E)_5\text{-}(C\text{-}A\text{-}PCP)_6\text{-}E_2$
M6a $(A\text{-}PCP)_6\text{-}E_2$	Fr-M6ATE-new.FOR AATCTTTATTTTCAGGCACAACCGCTTC CCATg Fr-M6ATE-new.REV GCAGCCGGATCAAGCTTACTTCTCGAAc	pET GB1 1d Tcp11 dimodule 5-6_E $(A\text{-}PCP\text{-}E)_5\text{-}(C\text{-}A\text{-}PCP)_6\text{-}E_2$	-
	V-M6ATE-new.REV CTGAAAATAAAGATTCTCAGAACCACT GCCAg	-	pET GB1 1d Tcp11 dimodule 5-

Final construct	**Primers**	**Insert template DNA**	**Vector template DNA**
	V-M6ATE-new.FOR GCTTGATCCGGCTGCTAACAAAG		6_E $(A\text{-}PCP\text{-}E)_5$- $(C\text{-}A\text{-}PCP)_6$- E_2
M7a $(C\text{-}A\text{-}PCP)_7COM_3$	F_Tcp12-CAT.FOR TATTTTCAGGGCGCCATGGGTACCGTG GATGACACGCGCGCTAAACCGCGCAGC TCAGTCGAGGACGTATGGCCGTTGTCT CCCCTTCAGGAGGGGATGCTGTATCAT ACCGCACTTG Fr-Tcp12-CAT.REV TTCACCATCCCCGCCCGATTTTGCTGCC AGAGCGC	pET MBP Tcp12 (CATXTe)	-
	V-COM.REV GGCGCCCTGAAAATAAAGATTCT V-COM.FOR GGCGGGGATGGTGAAGC	-	pET GB1 1 d Tcp10 (CAT)
M3a $C_7\text{-}(A\text{-}PCP)_3$	F_Tcp12- C-domain.FOR TATTTTCAGGGCGCCATGGGTACCGTG GATGACACGCGCGCTAAACCGCGCAGC TCAGTCGAGGACGTATGGCCGTTGTCT CCCCTTCAGGAGGGGATGCTGTATCAT ACCGCACTTG F-Tcp12_C-domain.REV ATCGACAGCATCACCAGTCGCGTTCCA GTCACCC	pET MBP Tcp12 (CATXTe)	-
	V_Tcp10.REV GGCGCCCTGAAAATAAAGATTCTC	-	pET GB1 1 d Tcp10 (CAT)

Final construct	**Primers**	**Insert template DNA**	**Vector template DNA**
	V-Tcp10.FOR GGTGATGCTGTCGATGCTC		
M4a (A-PCP-E-C)	FrTcp11 -Mod4 ATEC.FOR TATTTTCAGGGCGCCCCGGATCGCCTG GCGC FrTcp11 -Mod4 ATEC.REV GCAGCCGGATCAAGCTTACTTCTCGAA	pET GB1 1d Tcp11 module 4 (CATEC)	-
	Vector.REV GGCGCCCTGAAAATAAAGATTCTCAG Vector.FOR GCTTGATCCGGCTGCTAACAAAG	-	pET GB1 1d series
M5c C_4-(A-PCP-E-C)$_5$	Fragment 1.FOR AGAAGGAGATATACCATGAAACATCAC CAT Fragment 1.REV CTGCTCCCCGGACGTGCACCC	pET GB1 1d Tcp11 module 4 (CATEC)	-
	Fragment 2.FOR ACGTCCGGGGAGCAGCGTATTGACAG Fragment 2.REV CCAAGCACTCTCGAGCGTGAGCGGCGC ATCTGTTTCG	pET GB1 1d Tcp11 module 5 (ATEC)	-
	Vector.REV GGTATATCTCCTTCTTAAAGTTAAACAA AATTATTTCTAGAGGG Vector.FOR CTCGAGAGTGCTTGGAGTCATCC	-	pET GB1 1d Tcp11 module 4 (CATEC)
	Fr1-Mod4ATEC.FOR	pET GB1 1d Tcp11	-

Final construct	**Primers**	**Insert template DNA**	**Vector template DNA**
M4-6a (A-PCP-E)$_4$-(C-A-PCP-E)$_5$-(C-A-PCP)$_6$	AGAAGGAGATATACCATGAAACATCACCAT Fr1-Mod4ATEC.REV CTGCTCCCTCCTGCGGCCACGCC	module 4 (CATEC)	
	Fr2-56.FOR CGCAGGAGGGAGCAGCGTATTGACAGG Fr2-56.REV GCAGCCGGATCAAGCTTACTTCTCGAA	pET GB1 1d Tcp11 dimodule 5-6 (ATEC-AT)	
	Vector(Tcp11 FL).REV GGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAGGG Vector(Tcp11 FL).FOR GCTTGATCCGGCTGCTAACAAAG	-	pET GB1 1d series
M4d (C-A-PCP-**E-**$_5$C)	Fr1-M4CATEC.For(2) AGAAGGAGATATACCATGAAACATCACCATCACCATCACCCC Fr1-M4CATEC.REV(2) GCGTTGGCCCGCCGCCGGGGCCAACGCGGCCAGACGTTCTGGAGTACG	pET GB1 1d Tcp11 module 4 (CATEC)	-
	Fr2-Efrom5.FOR GGCGGCGGGCCAACGCCCAGCGGG Fr2-Efrom5.REV CAGCGGCCAGATATCCAGCAGGTCTGG	pET GB1 1d Tcp11 module 5 (ATEC)	-
	Vector.REV GGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAGGG Vector.FOR	-	pET GB1 1d Tcp11 module 4 (CATEC)

Final construct	**Primers**	**Insert template DNA**	**Vector template DNA**
	GATATCTGGCCGCTGAGCCC		
M4c (C-A-PCP-**E_5**-C)	F1-M4CAT.FOR AGAAGGAGATATACCATGAAACATCACCATC F1-M4CAT_REV CGTTGGCCCGCCGCCGGGGCCAACGCGGCCAG	pET GB1 1d Tcp11 module 4 (CATEC)	-
	F2-M5E.FOR GGCGGCGGGCCAACGCCCAGCGGG F2-M5E.REV GAAATCAGCCGGCGTATGGCCGCCTGCGC	pET GB1 1d Tcp11 module 5 (ATEC)	-
	V-M4COM.REV GGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAGGG V-M4COM.FOR ACGCCGGCTGATTTCGAC	-	pET GB1 1d Tcp11 module 4 (CATEC)
M4b (C-A-PCP-E-C) E-domain catalytic mutant (HHxxxDxxSW changed to HAxxxDxxSW)	Fr1_4CATEC(H-A).FOR AGGAGATATACCATGAAACATCACCATCACCA Fr1_4CATEC(H-A).REV CACCAGTGCGTGCGCGACCAGAGCC	pET GB1 1d Tcp11 module 4 (CATEC)	-
	Fr2_4CATEC(H-A).FOR GCGCACGCACTGGTGGTCGACGCAG Fr2_4CATEC(H-A).REV GCCGGATCAAGCTTACTTCTCGAA	pET GB1 1d Tcp11 module 4 (CATEC)	-
	Vector.REV	-	pET GB1 1d Tcp11

Final construct	**Primers**	**Insert template DNA**	**Vector template DNA**
	GGTATATCTCCTTCTTAAAGTTAAACAA AATTATTTCTAGAGGG Vector.FOR GCTTGATCCGGCTGCTAACAAAG		module 4 (CATEC)
M5a (A-PCP-E-C) E-domain catalytic mutant (HHxxxDxxSW changed to HQxxxDxxSW)	Fr1-Mod5Emut(H to Q).FOR AGGAGATATACCATGAAACATCACCAT CACCA Fr1-Mod5Emut(H to Q).REV GACCAGTTGGTGAGCTACAACCACTAA ACGT	pET GB1 1d Tcp11 module 5 (ATEC)	-
	Fr2-Mod5Emut(H to Q).FOR GCTCACCAACTGGTCGTTGACGCT Fr2 Mod5Emut(H to Q).REV GCCGGATCAAGCTTACTTCTCGAA	pET GB1 1d Tcp11 module 5 (ATEC)	-
	Vector.REV GGTATATCTCCTTCTTAAAGTTAAACAA AATTATTTCTAGAGGG Vector.FOR GCTTGATCCGGCTGCTAACAAAG	-	pET GB1 1d Tcp11 module 5 (ATEC)
M5b (A-PCP-E-C) E-domain catalytic mutant (HHxxxDxxSW changed to HAxxxDxxSW)	Fr1.Mod5Emut(H to A).FOR AGGAGATATACCATGAAACATCACCAT CACCA Fr1-Mod5Emut(H to A).REV GACCAGTGCGTGAGCTACAACCACTAA ACGT	pET GB1 1d Tcp11 module 5 (ATEC)	-
	Fr2-Mod5Emut(H to A).FOR GCTCACGCACTGGTCGTTGACGCT Fr2 Mod5Emut(H to A).REV	pET GB1 1d Tcp11	-

Final construct	**Primers**	**Insert template DNA**	**Vector template DNA**
	GCCGGATCAAGCTTACTTCTCGAA	module 5 (ATEC)	
	Vector.REV GGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAGGG Vector.FOR GCTTGATCCGGCTGCTAACAAAG	-	pET GB1 1d Tcp11 module 5 (ATEC)

Supplementary Table S2.

MS^2 analysis of each sample is provided in the separate excel sheet as an accompanying supplementary information file.

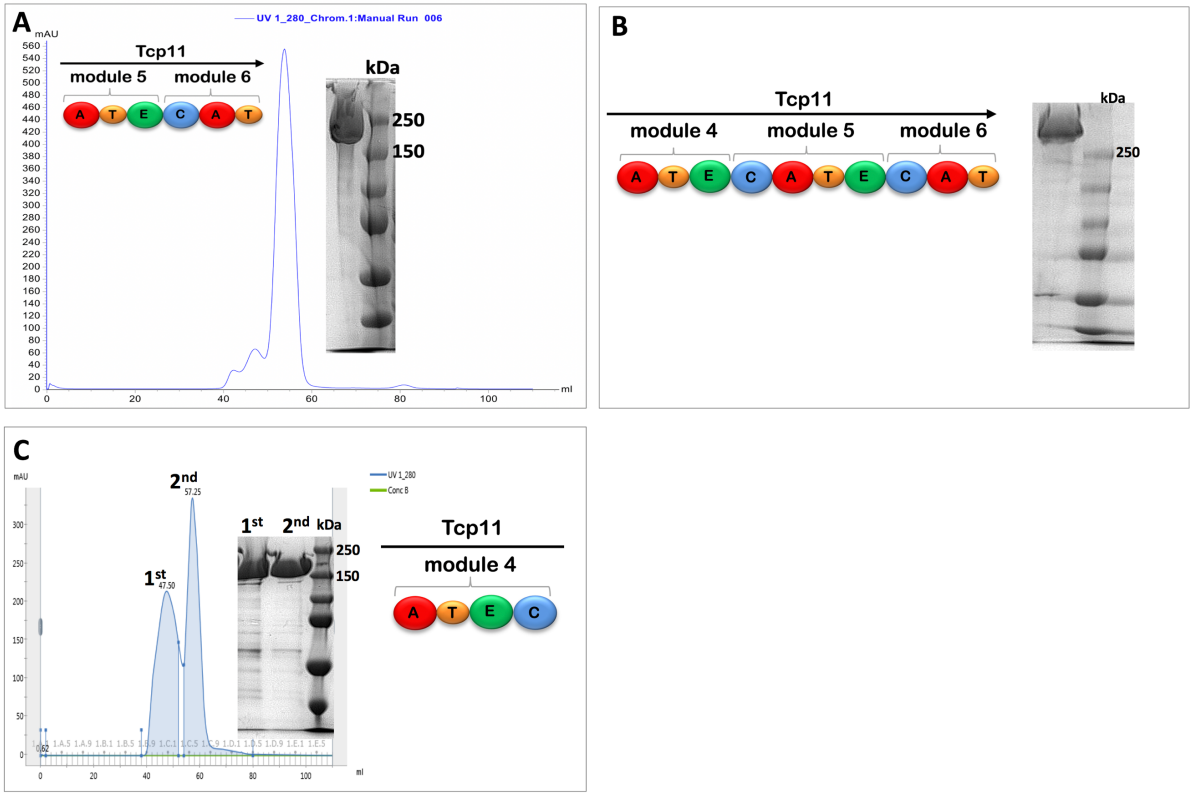
Detailed analysis for each detected peptide product in every reaction tested are listed in the table, where characteristic mass spectrometric parameters are listed, including *m/z*, peak height, retention time, putative sequence composition (NHMe-methylamide peptide, OH-hydrolysed peptide), mass error (ppm, theoretical vs detected), and the MS^2 ions detected.

Based on ions detected, putative sequences are provided using single letter code (C-ClBht, X-ClTyr, H-Hpg, D-Dpg, Y-Tyr) in the column "**Putative Sequence ID**". Based on our rationale identifying biosynthetic pathways, these putative sequences are divided into 4 groups using different colour code: (1) green – normal biosynthesis, where the sequence code that stems from the cyclic set of expected intermodule interactions is detected; (2) red – sequences that cannot be explained by the anticipated module interactions; (3) blue – sequences containing a partial sequence code that indicates the existence of unusual intermodule interactions; (4) pink – sequences that indicate an anticipated sequence that is undergone unusual N-terminal synthetic peptide extension or extensions at both C- and N-terminus the sequence. The column "**Sequence within NRPS Pathway**" indicates how these sequences fit within the anticipated peptide synthesis pathway.

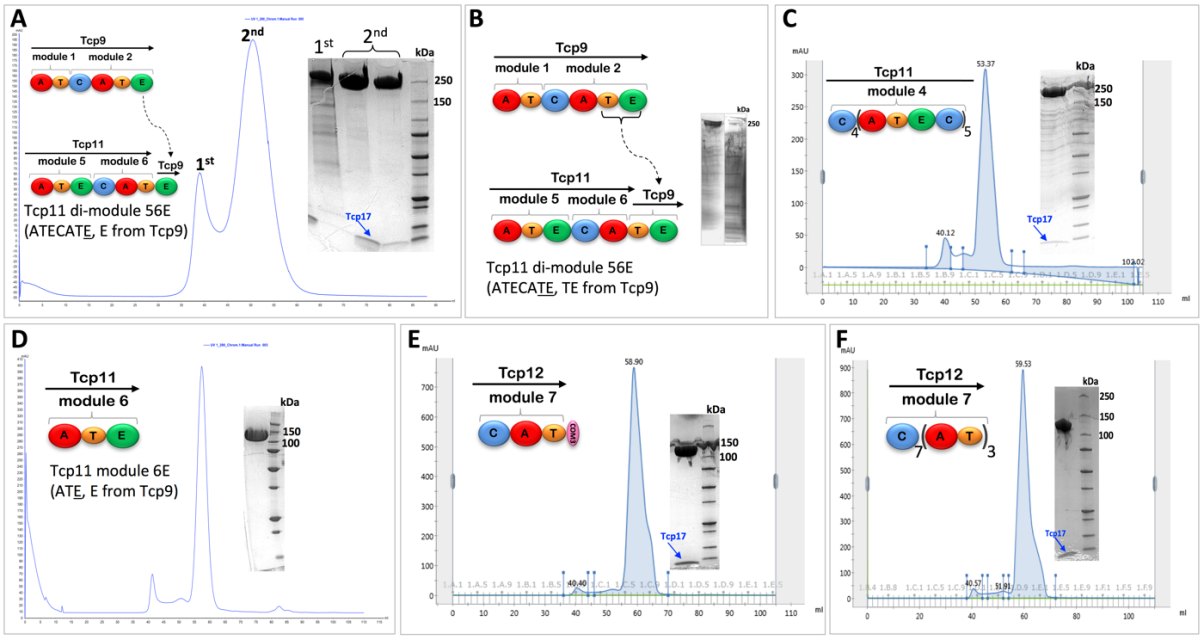
The proposed biosynthesis pathway leading to the formation of these sequences uses the following colour code shown in the "**Proposed Module Pathway**" column : (1) green – anticipated biosynthesis pathway; (2) orange – sequences that are in agreement with the anticipated biosynthesis pathway but are lacking complete MS^2 confirmation; (3) blue – sequence derived from unusual intermodule interactions. Grey in any column indicates pathway components that do not contribute to the sequence observed. Question marks indicate either uncertain sequences or difficulties in rationalising the formation of the specific peptide product.

Supplementary Table S3. Peptide species identified in each figure with their respective *m/z* values.

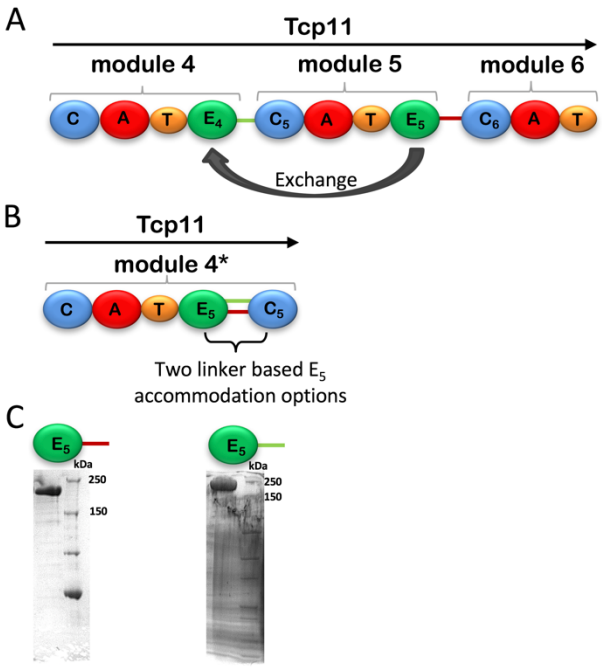
Peptide Species	*m/z* - LCMS	*m/z* - HRMS	Figure
1	363.1 (DKP)		**2, S11**
2	**2a** – 394.1 (NHMe) **2b** – 381.1 (OH)		**2, S11**
3	**3a** – 559.2 (NHMe) **3b** – 546.1 (OH)	**3a** – 559.15901	**2, 3**
4	299.1 (DKP)		**2**
5	**5a** – 330.1 (NHMe) **5b** – 317.1 (OH)		**2**
6	**6a** – 543.2 (NHMe) **6b** – 530.1 (OH)		**2**
7	**7a** – 708.2 (NHMe) **7b** – 695.2 (OH)	**7a** – 708.20669	**2, 6**
8	**8a** – 692.2 (NHMe) **8b** – 679.2 (OH)		**S9**
9	**9a** – 841.3 (NHMe) **9b** – 828.2 (OH)		**S9**
10	**10a** – 509.2 (NHMe) **10b** – 496.2 (OH)		**4, SI12**
11	**11a** – 658.2 (NHMe) **11b** – 645.2 (OH)	**11a** – 658.25075	**4, S12**
12	**12a** – 807.3 (NHMe) **12b** – 794.3 (OH)	**12a** – 807.29843	**4, S12**
13	**13a** – 708.2 (NHMe) **13b** – 695.2 (OH)	**13a** – 708.20669	**6, 7, S11**
14		**14a** – 857.25437	**6**
15		**15a** – 708.20669	**6, 7**
16		**16a** – 873.24928	**6**
17		**17a** – 857.25437	**6**
18		**18a** – 708.20669	**7**
19		**19a** – 1070.27364	**7**
20		**20a** – 873.24928	**7**
21		**21a** – 871.27002 **21b** – 858.23839	**8**
22		**22a** – 692.21178	**9**
23		**23a** – 841.25946	**9**
24		**24a** – 543.16410 **24b** – 530.13247	**9, S10**
25		**25b** – 844.22274	**9**
26		**26b** – 1041.24710	**9**
27		**27a** – 807.29843	**S12**
28		**28a** – 658.25075	**S12**
29		**29a** – 807.29843	**S12**



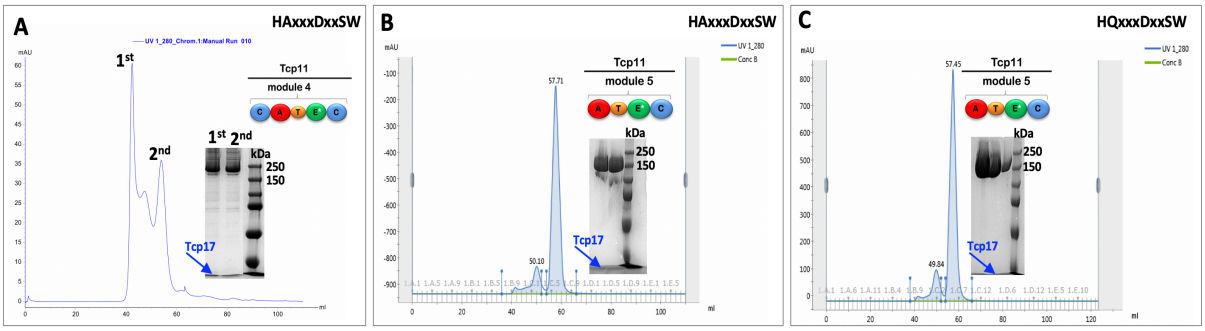
Supplementary Figure S1. Purification of alternate Tcp11 constructs. A) Gel filtration purification results of the **M5-6** construct (235 kDa). To improve solubility, this protein was fused with a GB1 solubility tag. B) Truncated version of the full length Tcp11 protein, **M4-6a** (GB1 solubility tag) after Strep-tag® purification (393 kDa); gel filtration chromatography was not performed. C) Gel filtration purification results of truncated version of module 4, **M4a** (GB1 solubility tag), 170 kDa. Only protein from the second fraction was collected.



Supplementary Figure S2. Protein purifications results and schematic protein architecture representation of four re-engineered NRPS modules. A) Gel filtration purification results of Tcp11 dimodule **M5-6a** (281 kDa, where E domain was relocated from Tcp9. Fractions from the second peak were used for biochemical characterisation. B) Tcp11 dimodule **M5-6b** (281 kDa), where the PCP and E didomain was relocated from Tcp9. Due to low protein yield only Ni-NTA and strep-tag purification stages were performed. C) Reengineered Tcp11 module 5, **M5c** (220 kDa), where the C-domain was from **M4**, A-PCP-E-C from **M5**; a single peak after gel filtration indicates correct protein folding. D) Re-engineered module 6, **M6a** (121 kDa), construct, where the E domain was relocated from Tcp9; a single peak after gel filtration indicates correct protein folding. E) Tcp12 truncated construct CAT_7COM_3, **M7a** (125 kDa), where X and Te domains were removed and short communication between **M3** and **M4** ensuring domain (COM_3) was fused with truncated construct at the protein C-terminus; a single peak after gel filtration indicates correct protein folding. F) Reengineered Tcp12 module 3, **M3a** where C domain was taken from **M7** and the A-PCP didomain from M3; a single peak after gel filtration indicates correct protein folding. To improve solubility, all proteins were fused with GB1 tag and were co-expressed with Tcp17 MbtH-like protein. Where it is visible, blue arrow indicates MbtH-like protein. T (thiolation) is another naming for a PCP domain.

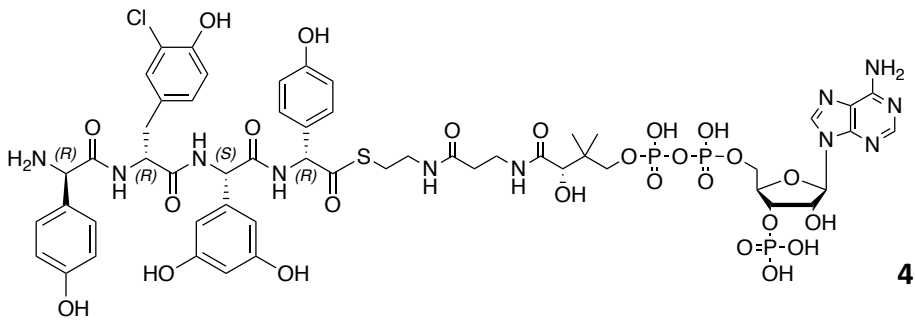


Supplementary Figure S3. Generation and purification of module 4 constructs with exchanged E-domains. Two linker regions between either E_4 and C_5 (shown in green) or between E_5 and C_6 (shown in red) were analysed (A and B) and corresponding constructs (219 kDa, GB1 fusions) were designed and purified (C). Due to low protein yield only Ni-NTA and strep-tag purifications were performed. T indicates a PCP domain. (C) **M4d** construct on the right and **M4c** on the left.



Supplementary Figure S4. Gel filtration purification results of both module 4 and module 5 E-domain catalytic mutants. Additional low molecular weight band indicate Tcp17 protein. To improve solubility all proteins were fused with GB1 tag. A) Module 4 E-domain active site mutant, **M4b** (HAxxxDxxSW), 218 kDa. Only the monomeric fraction from the second peak was used for biochemical characterisation. B) Module 5 E-domain active site mutant **M5b**, (HAxxxDxxSW), 171 kDa. C) Module 5 E-domain active site mutant **M5a**, (HQxxxDxxSW), 171 kDa. T indicates a PCP domain.

Supplementary Figure 5. Synthesis and characterisation of NH_2-(D)-Hpg-(D)-ClTyr-(L)-Dpg-(D)-Hpg-CoA tetrapeptide standard **(4*D*)-8-CoA**.



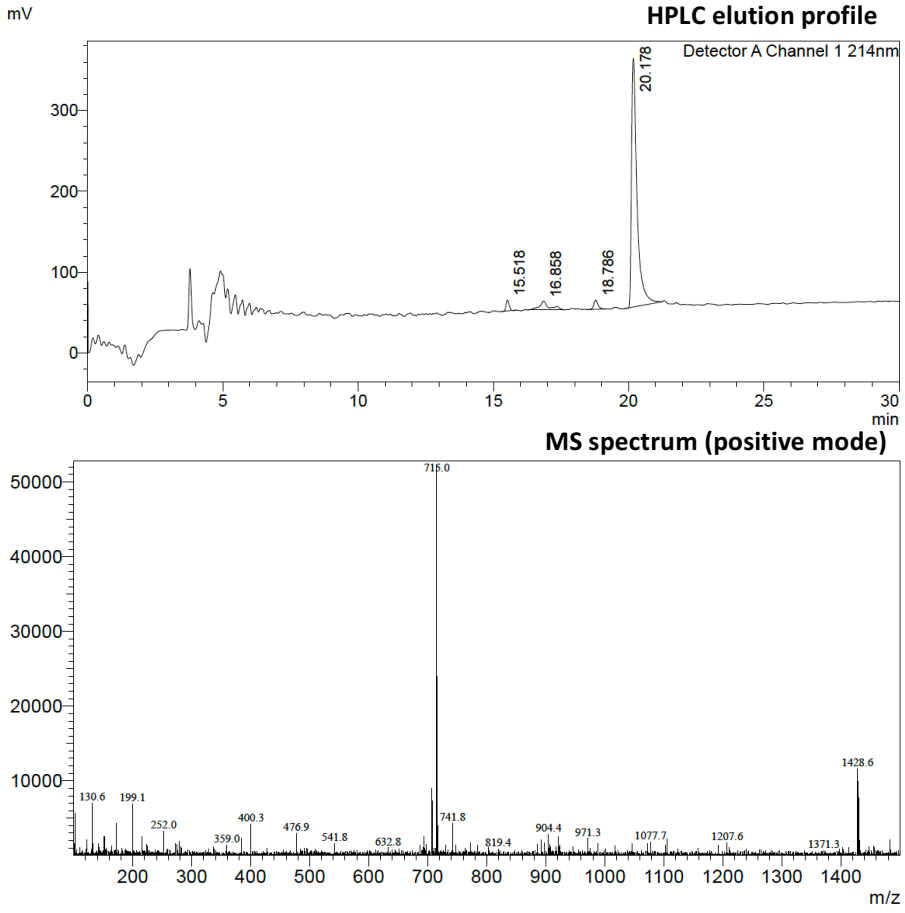
4 *D*-

Chemical Formula: $C_{54}H_{65}ClN_{11}O_{25}P_3S$

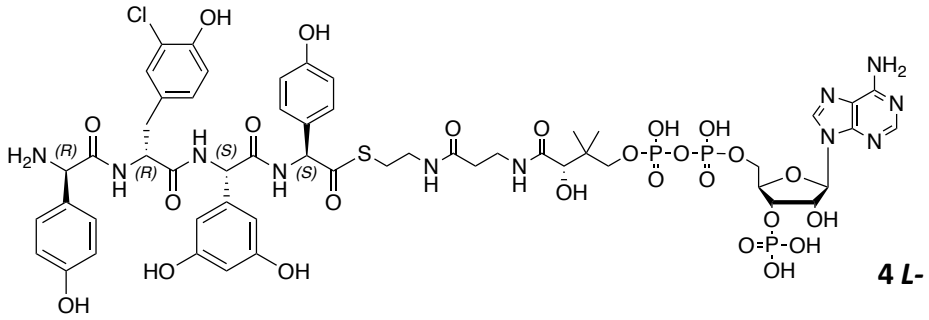
Exact Mass: 1427.28 Da

Retention time: 120.178 min (89%).

MS (ESI): *m/z* calcd $C_{54}H_{65}ClN_{11}O_{25}P_3S$ 1427.6, found *m/z* $C_{54}H_{66}ClN_{11}O_{25}P_3S^+$ 1428.6, $C_{54}H_{67}ClN_{11}O_{25}P_3S^{2+}$ 715.0.



Supplementary Figure 6. Synthesis and characterisation of NH_2-(D)-Hpg-(D)-ClTyr-(L)-Dpg-(L)-Hpg-CoA tetrapeptide standard **(4*L*)-8-CoA**.



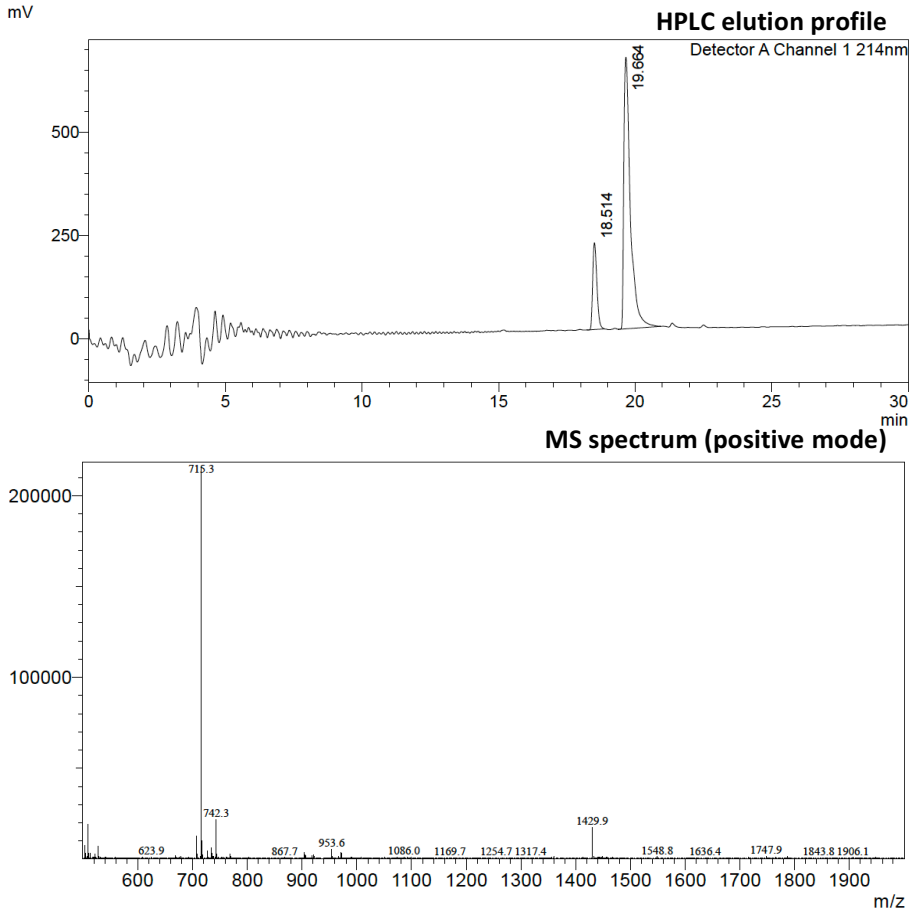
4 *L*-

Chemical Formula: $C_{54}H_{65}ClN_{11}O_{25}P_3S$

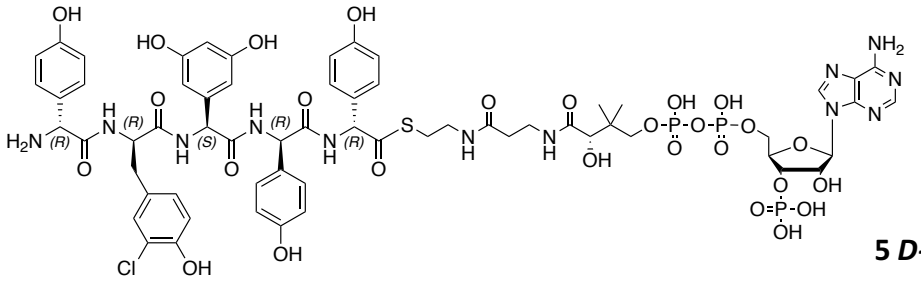
Exact Mass: 1427.28 Da

Retention time: 19.4664 min (83%).

MS (ESI): *m/z* calcd $C_{54}H_{65}ClN_{11}O_{25}P_3S$ 1428.6, found *m/z* $C_{54}H_{66}ClN_{11}O_{25}P_3S^+$ 1429.9, $C_{54}H_{67}ClN_{11}O_{25}P_3S^{2+}$ 715.3.



Supplementary Figure S7. Synthesis and characterisation of NH_2-(D)-Hpg-(D)-ClTyr-(L)-Dpg-(D)-Hpg-(D)-Hpg-CoA pentapeptide standard **(5*D*)-9-CoA**.



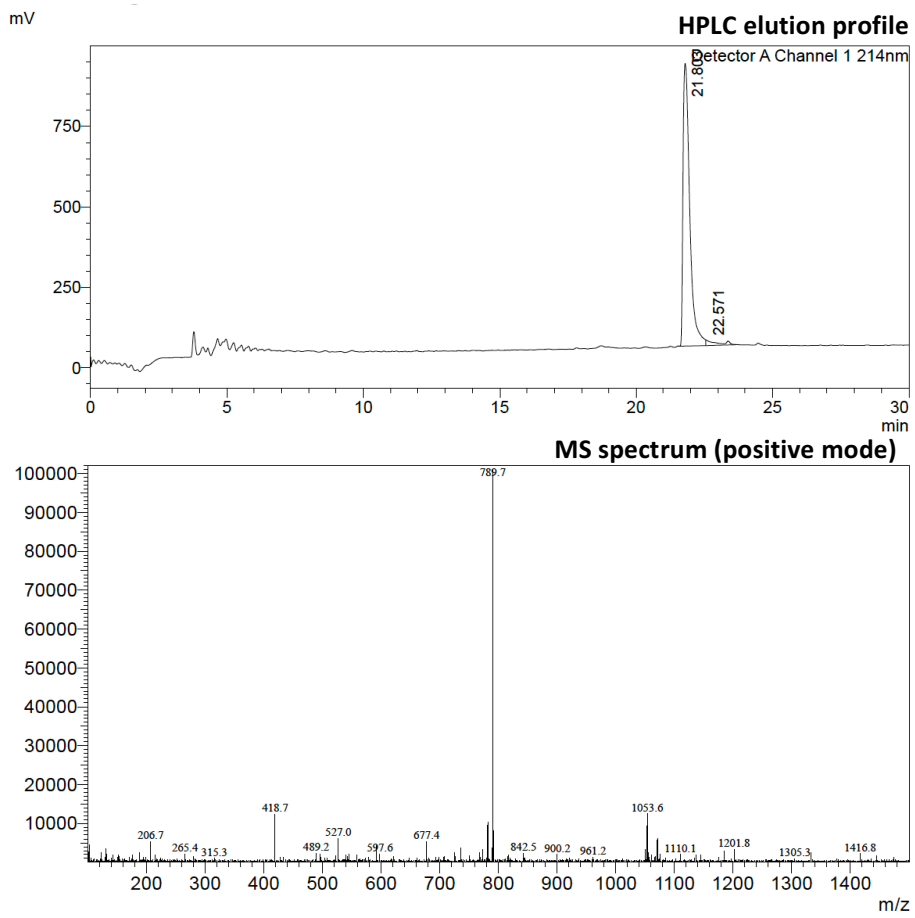
5 *D*-

Chemical Formula: $C_{62}H_{72}ClN_{12}O_{27}P_3S$

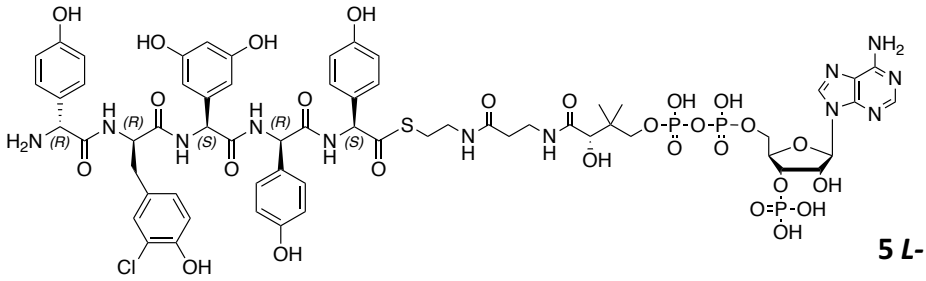
Exact Mass: 1576.3 Da

Retention time: 21.803 min (97%).

MS (ESI): *m/z* calcd $C_{62}H_{72}ClN_{12}O_{27}P_3S$ 1577.4, found *m/z* $C_{62}H_{74}ClN_{12}O_{27}P_3S^{2+}$ 789.7.



Supplementary Figure S8. Synthesis and characterisation of NH_2-(D)-Hpg-(D)-ClTyr-(L)-Dpg-(D)-Hpg-(L)-Hpg-CoA pentapeptide standard **(5*L*)-9-CoA**.



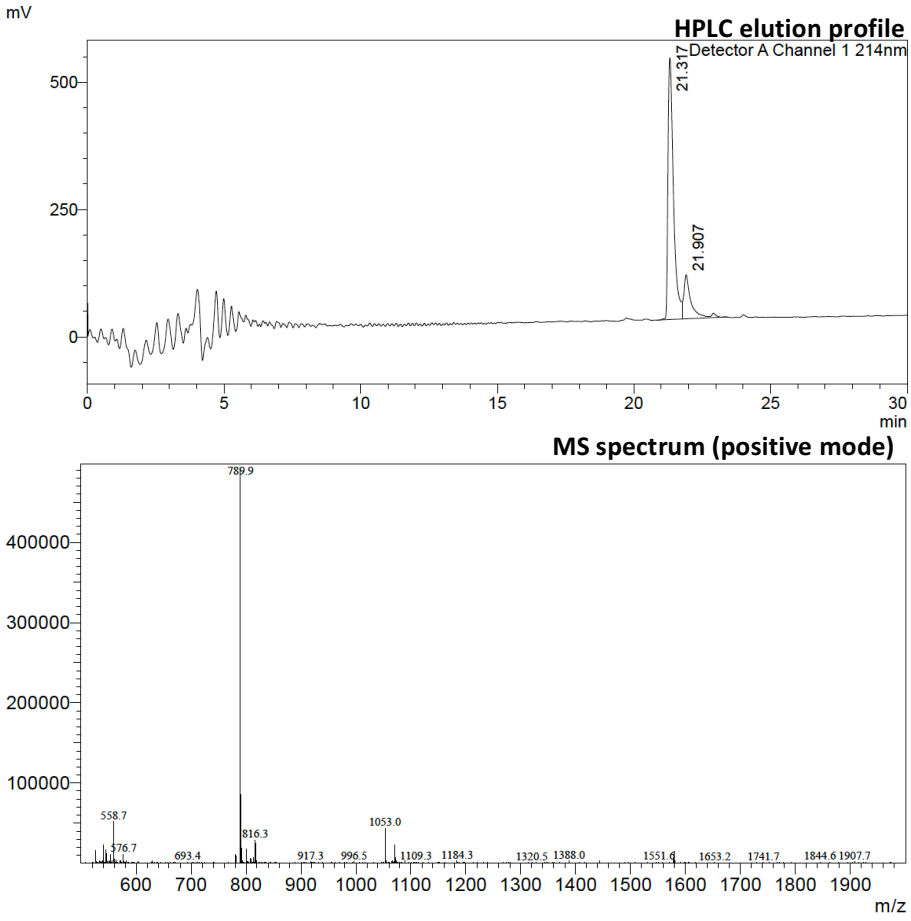
5 *L*-

Chemical Formula: $C_{62}H_{72}ClN_{12}O_{27}P_3S$

Exact Mass: 1578.3 Da

Retention time: 21.317 min (82%).

MS (ESI): *m/z* calcd $C_{62}H_{72}ClN_{12}O_{27}P_3S$ 1577.2, found *m/z* $C_{62}H_{74}ClN_{12}O_{27}P_3S^{2+}$ 789.9.



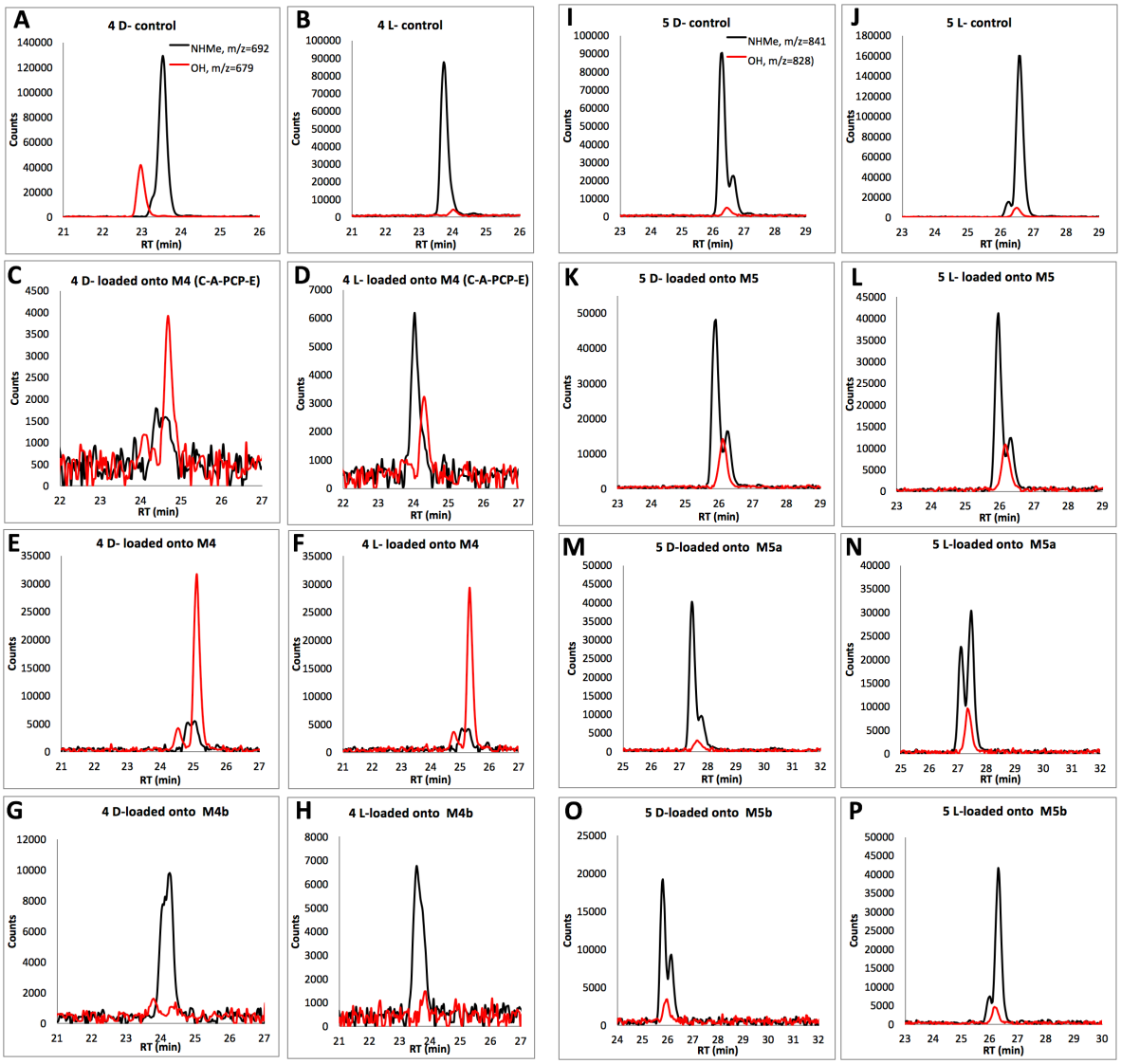


Figure S9. Summary of results of tetrapeptide and pentapeptide epimerisation analyses. A-B) LCMS traces showing **4*D*-8** and **4*L*-8** peptide standards (**SI Figures 5-6**) after 16h incubation in reconstitution assay buffer and cleavage with methylamine. C-D) LCMS traces showing the **4*D*-8** and **4*L*-8** peptides loaded onto the wild type **M4** (C-A-PCP-E architecture). E-F) LCMS traces showing **4*D*-8** and **4*L*-8** peptides loaded onto the alternate **M4** construct (C-A-PCP-E-C architecture). G-H) LCMS traces showing **4*D*-8** and **4*L*-8** peptides loaded onto **M4** with an inactive E-domain (**M4b** construct). I-J) LC-MS show **5*D*-9** and **5*L*-9** peptide standards (**SI Figures 7-8**) after 16h incubation in reconstitution assay buffer and cleavage with methylamine. K-L) LCMS traces showing **5*D*-9** and **5*L*-9** peptides loaded onto the wild type module 5 (**M5** construct). M-N) LCMS traces showing **5*D*-9** and **5*L*-9** peptides loaded onto **M5**

with an E-domain showing reduced activity (**M5a** construct). O-P) LCMS traces showing **5*D*-9** and **5*L*-9** peptides loaded onto the **M5** with an inactive E-domain (**M5b** construct).

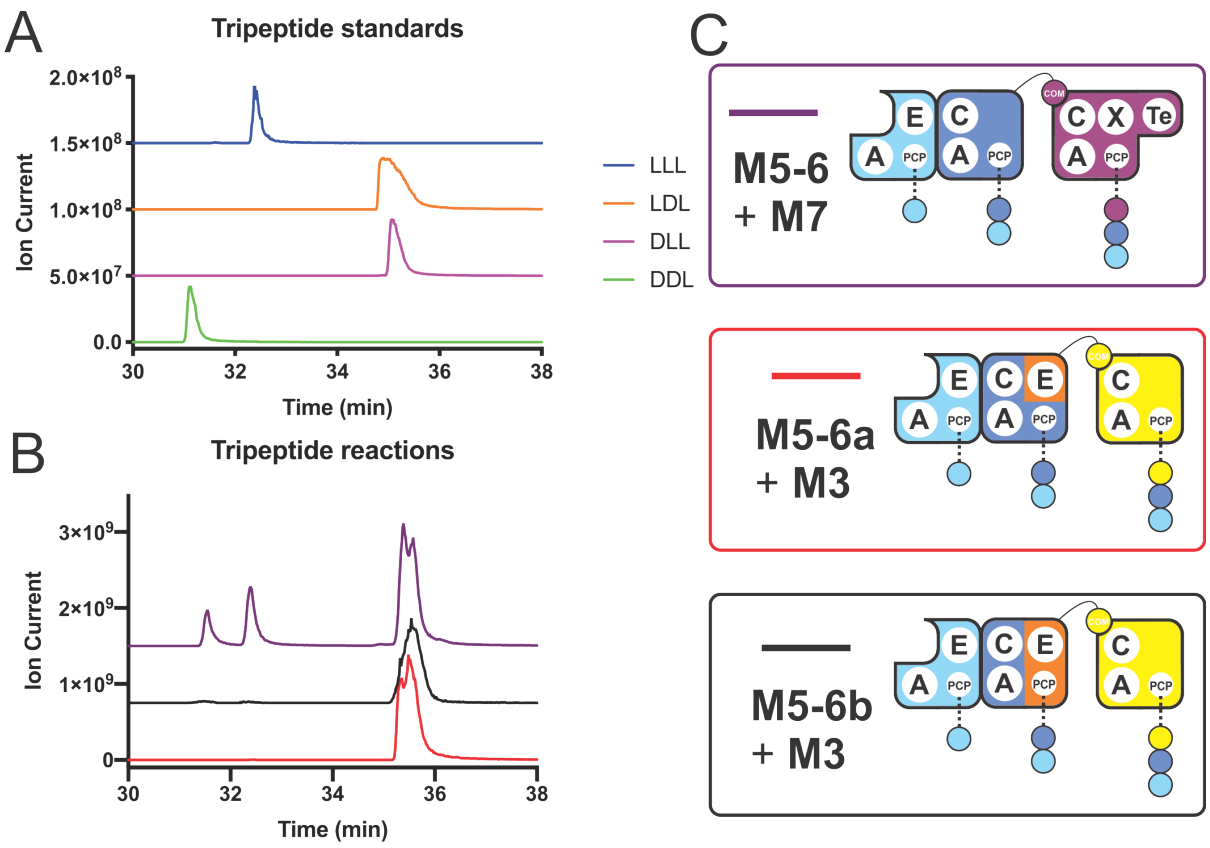


Figure S10. Analysis of the stereochemistry of Hpg-ClTyr-Dpg tripeptide 24 using LC-coupled HRMS. Comparison of the retention times of synthetic standards of **24** (reproduced from Kaniusaite *et al*,[1] stereochemistry shown for LLL (blue), DDL (green), LDL (orange) and DLL (pink) standards) (A). Tripeptide **24** produced by **M5-6** + **M7** (purple), **M5-6a** + **M3** (red) and **M5-6b** + **M3** (black) assembly lines together with ATP and the substrates of the A-domains of each module (M5 – Hpg; M6 – ClTyr; M3/M7 – 3,5-Dpg) (B). Schematic representation of NRPS assembly lines (C). For *m/z* data see **SI Table S3**. Domain key: A – adenylation, C – condensation, PCP – peptidyl carrier protein, E – epimerisation, X – P450-recruitment, TE – thioesterase, COM - communication. Module colour codes: M2 – orange, M3 – yellow, M5 – pale blue, M6 – dark blue, M7 – violet.

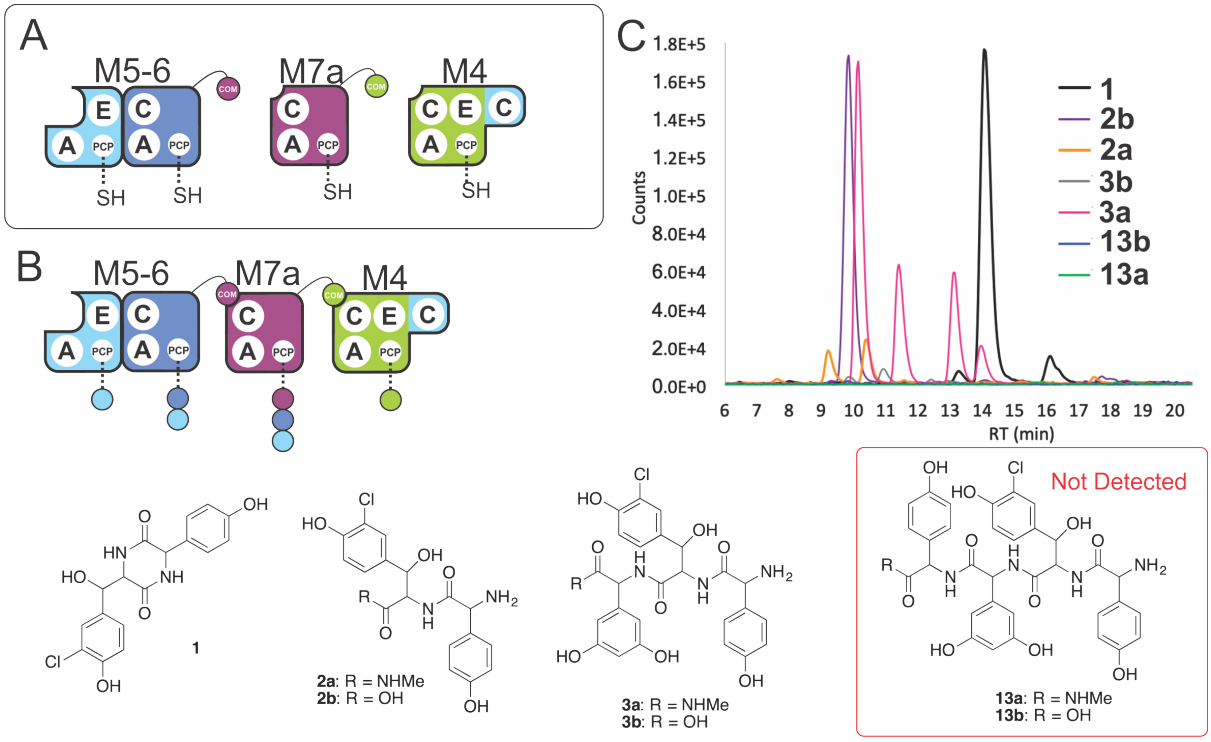


Figure S11. **COM-domain transplantation is insufficient to allow interaction across modules M7 and M4.** Peptide biosynthesis assays using dimodule **M5-6**, an engineered **M7** module **M7a** bearing the **M3** COM-domain and **M4** (A) together with ATP and the substrates of the A-domains of each module (M4, M5 – Hpg; M6 – ClBht; M7 – 3,5-Dpg) demonstrates that formation of tripeptide **3** is possible from **M5-6** + **M7**. However, extension of the assembly line with the addition of **M4** results in only formation of **3** (B) as indicated in LCMS traces (C). Peptide species indicated; for *m/z* data see **SI Table S3**. Domain key: A – adenylation, C – condensation, PCP – peptidyl carrier protein, E – epimerisation, COM - communication. Module colour codes: M4 – green, M5 – pale blue, M6 – dark blue, M7 – violet.

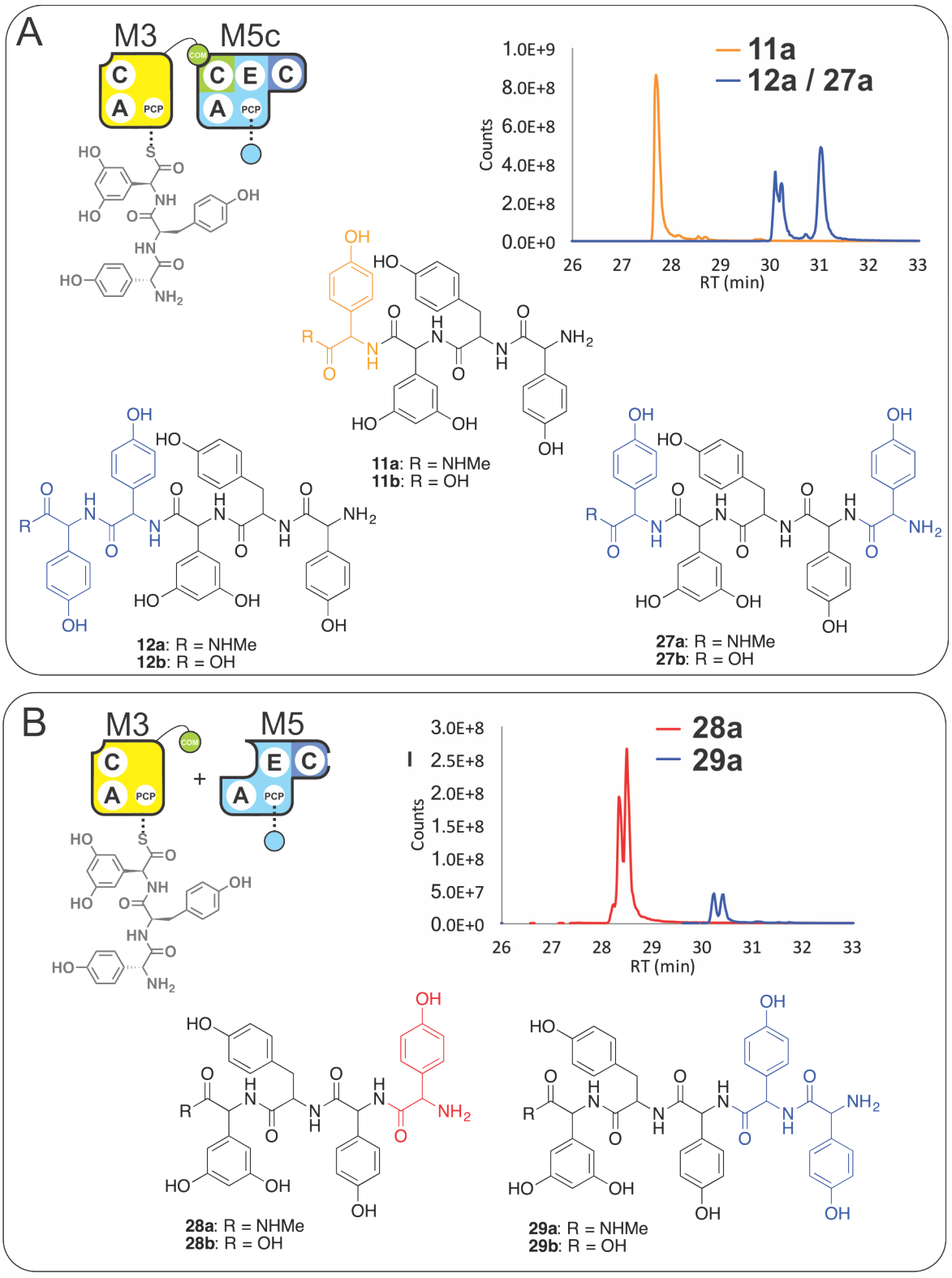


Figure S12. **Control reaction to test M3 communication with M5c/M5 constructs.** Peptide biosynthesis assays commencing from synthetic tripeptide **10** loaded on **M3** together with engineered **M5c** (A) or **M5** (B) together with ATP and the substrates of the A-domain of **M5** (Hpg). The **M3** + **M5c** reaction shows effective extension of **10** to **11** together with unexpected pentapeptide **12** & **27** , with MS/MS analysis indicating synthetic peptide elongation at the

peptide C- and N- termini (A). The **M3** + **M5** reaction shows unexpected tetrapeptide **28** and pentapeptide **27** production, with MS/MS analysis indicating the synthetic peptide was elongated at the N-terminus. Peptide species indicated; for *m/z* data see **SI Table S3**. Colour code indicates the amino acids added during the assays and matches the colours used in the LCMS traces. Domain key: A – adenylation, C – condensation, PCP – peptidyl carrier protein, E – epimerisation, COM - communication. Module colour codes: M3 – yellow, M4 – green, M5 – pale blue, M6 – dark blue.

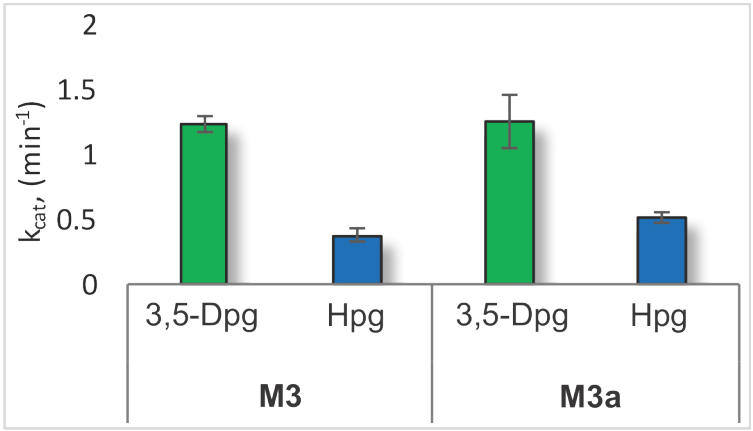


Figure S13. **Exchanging the C-domain within M3 constructs does not alter A-domain specificity in these constructs.** A-domain activity assays performed with **M3** and **M3a** using their specific (3,5-Dpg) and alternative (Hpg) amino acid substrates. A-domain activity measurements are generated using a coupled enzyme activity assay[2] performed in triplicate with the standard deviation shown.

Supplementary References

1. M. Kaniusaite, J. Tailhades, T. Kittilä, C. D. Fage, R. J. A. Goode, R. B. Schittenhelm and M. J. Cryle, *The FEBS Journal*, 2020, DOI: 10.1111/febs.15350, 10.1111/febs.15350.
2. T. Kittilä, M. Schoppet and M. J. Cryle, *ChemBioChem*, 2016, **17**, 576-584.